

RECOVERING FOURIER COEFFICIENTS OF MODULAR FORMS AND FACTORING OF INTEGERS

Sergei N. Preobrazhenskii¹

It is shown that if a function defined on the segment $[-1, 1]$ has sufficiently good approximation by partial sums of the Legendre polynomial expansion, then, given the function's Fourier coefficients c_n for some subset of $n \in [n_1, n_2]$, one can approximately recover them for all $n \in [n_1, n_2]$. As an application, a new approach to factoring of integers is given.

Key words: computational number theory, complexity of computing, algorithm, factorization, factoring of integers, elliptic curves, modular forms, Fourier coefficients, Legendre polynomials.

1. Introduction. Let a function $f(x)$ being a polynomial of degree K in the interval $[-\frac{1}{2}, \frac{1}{2})$ be continued periodically to the entire real axis with period 1:

$$f(x) = b_0 + b_1x + \dots + b_Kx^K, \quad x \in \left[-\frac{1}{2}, \frac{1}{2}\right), \quad b_K \neq 0, \quad f(x+1) = f(x).$$

A question is this: Can we find all Fourier coefficients of the function $f(x)$ if we know $K+1$ values of the coefficients c_{p_0}, \dots, c_{p_K} ? Obviously, one may set up the system of linear equations

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,K+1} \\ a_{2,1} & \dots & a_{2,K+1} \\ \dots & \dots & \dots \\ a_{K+1,1} & \dots & a_{K+1,K+1} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_K \end{pmatrix} = \begin{pmatrix} c_{p_0} \\ c_{p_1} \\ \vdots \\ c_{p_K} \end{pmatrix}.$$

So if we prove that the matrix of the system is nonsingular then we can uniquely identify the polynomial and find the remaining Fourier coefficients of the function $f(x)$.

In this note we show that if a function defined on the segment $[-1, 1]$ has sufficiently good approximation by partial sums of the Legendre polynomial expansion, then, given the function's Fourier coefficients c_n for some subset of $n \in [n_1, n_2]$, one can approximately recover them for all $n \in [n_1, n_2]$. As an application, a new approach to factoring of integers is given. Among factoring algorithms now in use the fastest ones are: the elliptic curve method, the quadratic sieve and the number field sieve method. The quadratic sieve algorithm was introduced by Pomerance [1, 2] in 1981. The heuristic complexity of the algorithm is

$$\exp((1+o(1))(\log n)^{1/2}(\log \log n)^{1/2})$$

arithmetic operations, where n is the number to be factored. The elliptic curve method of H. Lenstra [3, 4] appeared in 1986. If p denotes the least prime factor of n , then the expected number of operations required to factor n with the elliptic curve method is

$$\exp((\sqrt{2}+o(1))(\log p)^{1/2}(\log \log p)^{1/2})(\log n)^{C_2}, \quad (1)$$

where C_2 is a positive constant. The number field sieve was suggested in [5]. The heuristic complexity for factoring n via this method is

$$\exp((C_1+o(1))(\log n)^{1/3}(\log \log n)^{2/3}).$$

¹*Preobrazhenskii Sergei Nikolayevich* — Department of Mathematical Analysis, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University.

The main idea underlying the approach suggested in this note is to compute Fourier coefficients of modular forms that arise from elliptic curves (see [6, 7]).

Here is an outline of the method. It is known that certain modular forms have integer Fourier coefficients c_n , and these coefficients satisfy specific equations (see below). Moreover, the famous Shimura–Taniyama conjecture (proved by Wiles for the case of semistable elliptic curves) states that under certain conditions the numbers a_p — defined in terms of the orders over \mathbb{Z}_p of a fixed elliptic curve E over \mathbb{Q} — coincide with the Fourier coefficients c_p of a modular form of the aforementioned type. The numbers a_p are defined by the equation: $\#E(\mathbb{Z}_p) = p + 1 - a_p$. One may use these properties to factor $n = pq$. Let p and q be odd primes of about the same size (such n 's are called RSA-numbers). Choose arbitrary elliptic curve E over \mathbb{Q} . If for the corresponding modular form we could find the coefficients c_n and c_{n^2} , factor them, and thereby find, say, c_p and c_{p^2} , then from the aforementioned equations for the coefficients we would find $p = (c_p)^2 - c_{p^2}$. The problem is to find c_n and c_{n^2} when the factorization $n = pq$ is unknown. (If the factorization was known, we would find c_n and c_{n^2} by computing $c_p = a_p$ and $c_q = a_q$ via the Schoof algorithm for the curve E .) Thus, we wish to reduce the problem of factoring n to the problem of factoring c_n and c_{n^2} that depend on a randomly chosen curve E .

Assuming that the modular form on the interval $[-1, 1]$ has sufficiently good approximation by partial sums of the Legendre polynomial expansion, one could try to find c_n and c_{n^2} , factorization of n being unknown, via the “approximation” method as it is enunciated at the beginning of the section. One represents Fourier coefficients of the modular form as integrals, uses the polynomial approximation and the Schoof algorithm for computing small sets of coefficients $c_{p'}$ and $c_{p''}$ for small sets of primes $\{p'\}$ and $\{p''\}$, near n and n^2 , respectively. The coefficients in the “knots”, the sets $\{p'\}$ and $\{p''\}$, “interpolate” the coefficients in the “points” n and n^2 .

2. Main theorem. Suppose $\tau = \rho + i\sigma$, $\sigma = \frac{1}{n}$ is fixed, $\rho \in [-1, 1]$. Write the function $\varphi(\tau) = f(\tau)e^{-2\pi i n \tau}$ as the sum of its real and imaginary part:

$$\varphi(\tau) = f(\tau)e^{-2\pi i n \tau} = u_\sigma(\rho) + iv_\sigma(\rho).$$

Suppose

$$n < p'_1 = n + \Delta_1 < p'_2 = n + \Delta_2 < \dots < p'_K = n + \Delta_K.$$

For an integer Δ , denote by $c_\Delta(f)$ the Δ -th Fourier coefficient of $\varphi(\tau)$. It coincides with $(n + \Delta)$ -th coefficient of $f(\tau)$:

$$2c_\Delta(f) = e^{2\pi\Delta/n} \int_{-1}^1 (u_\sigma(\rho) + iv_\sigma(\rho)) e^{-i(2\pi\Delta)\rho} d\rho.$$

Take the finite Legendre polynomial expansion for the real and the imaginary part of $\varphi(\tau)$:

$$\begin{aligned} u_\sigma(\rho) &\sim \alpha_1 P_1(\rho) + \dots + \alpha_K P_K(\rho) = U_K(\rho), \\ v_\sigma(\rho) &\sim \beta_1 P_1(\rho) + \dots + \beta_K P_K(\rho) = V_K(\rho). \end{aligned} \tag{2}$$

Denote by $C_\Delta(\bar{\alpha}, \bar{\beta})$ the approximation to the Fourier coefficient $c_\Delta(f)$ obtained via the expansion (2):

$$2C_\Delta(\bar{\alpha}, \bar{\beta}) = e^{2\pi\Delta/n} \int_{-1}^1 (U_K(\rho) + iV_K(\rho)) e^{-i(2\pi\Delta)\rho} d\rho.$$

and

$$|2C_{\Delta_0}(\bar{\alpha} + \bar{\delta}', \bar{\beta} + \bar{\delta}'') - 2c_{\Delta_0}(f)| < E. \quad (7)$$

So if we know the Fourier coefficients $2c_{\Delta_1}(f), \dots, 2c_{\Delta_K}(f)$, then we may recover the Fourier coefficient $2c_{\Delta_0}(f)$ ($\Delta_0 \neq 0$) to the precision E .

Remark. In this theorem, instead of using Legendre polynomials, one could use the simple powers x, x^2, \dots, x^K . But in practice, we cannot know that (5) holds, and with Legendre polynomials we may infer this from stabilization of the coefficients

$$\begin{pmatrix} \alpha_1 + i\beta_1 + \delta'_1 + i\delta''_1 \\ \vdots \\ \alpha_K + i\beta_K + \delta'_K + i\delta''_K \end{pmatrix}$$

in (6) as K grows. For simple powers such stabilization of coefficients may not occur.

Proof. We now prove (4). Using the notation (3) and known Hankel expansions for Bessel functions of half-integer order, we deduce that $a_{k,m}(\Delta_k)$ are polynomials in $\frac{1}{\Delta_k}$ of degree m with the leading coefficients

$$b_{k,m} = \begin{cases} \frac{(-i)^m}{\pi} \cos\left(-\frac{(m+1/2)\pi}{2} - \frac{\pi}{4}\right) (-1)^{\mu-1} \frac{(m+1/2, m-1)}{(4\pi)^{m-1}}, & \text{if } m = 2\mu - 1; \\ \frac{(-i)^m}{\pi} (-1) \sin\left(-\frac{(m+1/2)\pi}{2} - \frac{\pi}{4}\right) (-1)^{\mu-1} \frac{(m+1/2, m-1)}{(4\pi)^{m-1}}, & \text{if } m = 2\mu. \end{cases}$$

By elementary operations, the matrix $\mathbf{A}_{K \times K}$ may be transformed to the Vandermonde matrix. This completes the proof of (4).

We now prove that the condition (5) implies (6) and (7). Define

$$\varphi_\sigma(\rho) = u_\sigma(\rho) + iv_\sigma(\rho), \quad \Phi_K(\rho) = U_K(\rho) + iV_K(\rho).$$

We have

$$\begin{aligned} \int_{-1}^1 |\varphi_\sigma(\rho) - \Phi_K(\rho)|^2 d\rho &= \int_{-1}^1 (\varphi_\sigma(\rho) - \Phi_K(\rho)) (\overline{\varphi_\sigma(\rho)} - \overline{\Phi_K(\rho)}) d\rho = \\ &= \int_{-1}^1 |\varphi_\sigma(\rho)|^2 d\rho - \int_{-1}^1 \Phi_K(\rho) \overline{\varphi_\sigma(\rho)} d\rho - \int_{-1}^1 \overline{\Phi_K(\rho)} \varphi_\sigma(\rho) d\rho + \int_{-1}^1 |\Phi_K(\rho)|^2 d\rho = \\ &= \int_{-1}^1 |\varphi_\sigma(\rho)|^2 d\rho - \sum_{k=1}^K \frac{2}{2k+1} (\alpha_k^2 + \beta_k^2) = \sum_{k=K+1}^{\infty} \frac{2}{2k+1} (\alpha_k^2 + \beta_k^2) = R_K(f). \end{aligned}$$

Let us estimate $|2C_{\Delta_k} - 2c_{\Delta_k}(f)|$. Using the Cauchy inequality and the condition (5), for $1 \leq k \leq K$ we obtain the estimates

$$\begin{aligned} |2C_{\Delta_k} - 2c_{\Delta_k}(f)| &\leq e^{2\pi\Delta_k/n} \int_{-1}^1 |\varphi_\sigma(\rho) - \Phi_K(\rho)| d\rho \leq \\ &\leq e^{2\pi\Delta_k/n} \sqrt{2} \left(\int_{-1}^1 |\varphi_\sigma(\rho) - \Phi_K(\rho)|^2 d\rho \right)^{1/2} < D_{K,\Delta_0}(E), \end{aligned}$$

from which by the definition of $D_{K,\Delta_0}(E)$ we infer (6) and (7). This completes the proof of the theorem.

3. Modular forms and modular elliptic curves. We follow the book [8].

An unrestricted modular form of level N and weight 2 is an analytic function f on H such that for all $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$ and $\tau \in H$ we have

$$f\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^2 f(\tau).$$

Since $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_0(N)$ then $f(\tau + 1) = f(\tau)$ for each modular form and every τ . Thus f has a Fourier expansion, which is of the form

$$f(\tau) = \sum_{n=0}^{\infty} c_n e^{2\pi i n \tau}. \quad (8)$$

Here

$$c_n = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\tau) e^{-2\pi i n \tau} d\rho,$$

where $\tau = \rho + i\sigma$, and $\sigma > 0$ is fixed. For cusp forms, $c_0 = 0$.

The following theorem conveys information about the magnitude of the modulus of a cusp form $f(\tau)$ and the Fourier coefficients c_n .

Theorem (Deuring). *Let $f \in \mathcal{S}_2(N)$ have q -expansion (8) at the cusp ∞ . Then the function $\varphi(\tau) = |f(\tau)|\sigma$ is bounded on H and invariant under $\Gamma_0(N)$. Furthermore, we have $|c_n| \leq Cn$.*

Suppose $f \in \mathcal{S}_2(N)$ is an eigenform, normalized so that the q expansion (8) has $c_1 = 1$. Then for $r \geq 1$ the Fourier coefficients c_n of f satisfy

$$\begin{aligned} c_{p^r} c_p &= c_{p^{r+1}} + p c_{p^{r-1}}, & \text{for } p \text{ prime, } p \nmid N; \\ c_{p^r} &= (c_p)^r, & \text{for } p \text{ prime, } p \mid N; \\ c_m c_n &= c_{mn}, & \gcd(m, n) = 1. \end{aligned} \quad (9)$$

The following statement is the Shimura–Taniyama conjecture, proved by Wiles for the case of semistable elliptic curves (i.e. for squarefree N).

Proposition. *If E is any elliptic curve defined over \mathbb{Q} , if N is its conductor, then there is a normalized new cusp eigenform f of level N and weight 2, whose Fourier coefficients c_n are integers and such that for every prime p not dividing N $c_p = a_p$ (where a_p is defined by $\#E(\mathbb{F}_p) = p + 1 - a_p$, $\#E(\mathbb{F}_p)$ being the order of the group of the elliptic curve E over \mathbb{F}_p).*

Now we give an example of a modular form associated with an elliptic curve.

Example (Hecke). Define

$$\Delta(\tau) = (2\pi)^{12} q \prod_{n=1}^{\infty} (1 - q^n)^{24} \quad \text{and} \quad \eta(\tau) = \frac{1}{\sqrt{2\pi}} \Delta(\tau)^{\frac{1}{24}}.$$

Then it follows that $f(\tau) = \eta(11\tau)^2 \eta(\tau)^2$ is a new cusp eigenform of weight 2 and level 11. It is associated with the elliptic curve E

$$y^2 + y = x^3 - x^2 - 10x - 20,$$

which for $p \neq 2, 3$ can be given the form

$$y^2 = x^3 - \frac{31}{3}x - \frac{2431}{108}.$$

4. Elliptic curves over finite fields. The following theorem gives a range for the order of an elliptic curve group defined over a finite field.

Theorem (Hasse). *Let E be an elliptic curve over \mathbb{Q} with integer coefficients and the discriminant Δ . For each prime $p \nmid \Delta$, let $E(\mathbb{F}_p)$ be the reduction of E modulo p . Then $|p + 1 - \#E(\mathbb{F}_p)| < 2\sqrt{p}$.*

We need two celebrated algorithms related to elliptic curves over finite fields: the Schoof algorithm [9] for elliptic curve point-counting and Lenstra's elliptic curve method for factorization, mentioned in Section 1. For prime $q > 3$ and elliptic curve $E(\mathbb{F}_q)$, the Schoof algorithm computes $\#E(\mathbb{F}_q)$ in complexity $O(\log^8 q)$. The Lenstra method factors a composite n . If p denotes the least prime factor of n , then this method has the complexity (1).

5. Smooth numbers. A positive integer is said to be y -smooth if it does not have any prime factor exceeding y . Let $\psi(x, y)$ denote the number of y -smooth integers among positive integers $n \leq x$.

Theorem (Canfield–Erdős–Pomerance [10]). *The estimate $\psi(x, x^{1/u}) = xu^{-u+o(u)}$ holds uniformly as $u \rightarrow \infty$, $u < (1 - \varepsilon_1) \log x / \log \log x$, $\varepsilon_1 > 0$.*

6. The central result and the algorithm.

Theorem 2. *Assuming that the method discussed in Section 2 for computing the Fourier coefficients of a modular form associated with an elliptic curve is correct and runs in polynomial time, the heuristic complexity estimate for factoring an integer n with the least prime factor p ($p^2 \nmid n$) is*

$$\exp(c_1(k)(\log p)^{1/k}(\log \log p)^{1-1/k}) (\log n)^{c_2(k)}$$

arithmetic operations for any fixed positive integer k ; here $c_1(k)$, $c_2(k)$ are positive constants which may depend on the choice of k .

We now describe a recursive algorithm having the claimed complexity estimate. First, we show that the following method gives the complexity estimate

$$\exp(c_1(3)(\log p)^{1/3}(\log \log p)^{2/3}) (\log n)^{c_2(3)}$$

(this method corresponds to the case $k = 3$).

Step 0. Choose a smoothness parameter $B \asymp p^{1/((\log p)^{1/3}(\log \log p)^{-1/3})}$, where p is the least prime factor of n to be factored. (In fact, since we do not know p to begin with, we are instructed to start with a low B_0 value and then to run the algorithm with $B = B_0$, $B = 2B_0$, $B = 4B_0$, \dots , choosing for each B about

$$\exp((\sqrt{2} + o(1))(\log B)^{1/2}(\log \log B)^{1/2})$$

random elliptic curves in Step 1.)

Step 1. Choose random elliptic curve E over \mathbb{Q} with integer coefficients.

Step 2. Compute the Fourier coefficients c_n and c_{n^2} of the associated modular form with the method discussed in Section 2.

Step 3. Try to factor these coefficients with the elliptic curve method, hoping for a successful event: the largest prime factor of the coefficient $c_p \mid c_n$ and the largest prime factor of the

coefficient $c_{p^2} \mid c_{n^2}$ both do not exceed B . For each $c_1 \mid c_n$ and $c_2 \mid c_{n^2}$ compute $d = \gcd((c_1)^2 - c_2, n)$ and, if $1 < d < n$, return the nontrivial divisor d of n .

Step 4. Failure: goto Step 1 (or give up on the factorization attempt).

Assume that the Fourier coefficients c_p and c_{p^2} of the modular form associated with the chosen elliptic curve (where $c_p = O(\sqrt{p})$ and $c_{p^2} = O(p)$ by the Hasse theorem and the relation (9)) are B -smooth with the same probability as random integers chosen from the respective intervals. Then we get B -smooth coefficients c_p and c_{p^2} expecting about $u^{\frac{3}{2}}$, where $u = (\log p)^{1/3}(\log \log p)^{-1/3}$, elliptic curves, with the attempt to find these coefficients in Step 3 requiring

$$\exp \left((\sqrt{2} + o(1))(\log B)^{1/2}(\log \log B)^{1/2} \right) (\log n)^{c'_2}$$

arithmetic operations. Multiplying the expressions, we obtain the estimate for the complexity of the algorithm in the form

$$\exp \left(c_1(3)(\log p)^{1/3}(\log \log p)^{2/3} \right) (\log n)^{c_2(3)}.$$

If we now use this algorithm instead of the elliptic curve method in Step 3 and choose the optimal smoothness parameter, then we get an algorithm having complexity

$$\exp \left(c_1(4)(\log p)^{1/4}(\log \log p)^{3/4} \right) (\log n)^{c_2(4)}.$$

Furthermore, if we use $k + 1$ levels of recursion, choose $B \asymp p^{1/((\log p)^{1/(k+1)}(\log \log p)^{-1/(k+1)})}$ and apply the algorithm with k levels of recursion, having complexity

$$\exp \left(c_1(k)(\log p)^{1/k}(\log \log p)^{1-1/k} \right) (\log n)^{c_2(k)},$$

then we arrive at the following complexity estimate:

$$\begin{aligned} & \left((\log p)^{1/(k+1)}(\log \log p)^{-1/(k+1)} \right)^{\frac{3}{2}(\log p)^{1/(k+1)}(\log \log p)^{-1/(k+1)}} \times \\ & \times \exp \left(c(k)(\log B)^{1/k}(\log \log B)^{1-1/k} \right) (\log n)^{c'_2(k)} = \\ & = \exp \left(c'_1(k)(\log p)^{1/(k+1)}(\log \log p)^{1-1/(k+1)} \right) \times \\ & \times \exp \left(c''_1(k)(\log p)^{(1/k)(1-1/(k+1))}(\log \log p)^{1/(k(k+1))+1-1/k} \right) (\log n)^{c'_2(k)} = \\ & = \exp \left(c_1((k+1))(\log p)^{1/(k+1)}(\log \log p)^{1-1/(k+1)} \right) (\log n)^{c_2((k+1))}. \end{aligned}$$

References

1. *Pomerance C.* Analysis and comparison of some integer factoring algorithms // Computational methods in number theory. V.1 / Ed. by H.W. Lenstra and R. Tijdeman. Amsterdam, 1982. 89–139.
2. *Pomerance C.* The quadratic sieve factoring algorithm // Advances in cryptology: Proc. Conf. (Paris, 1984). Lect. Notes Comput. Sci. 1985. **209**. 169–183.
3. *Lenstra H. W.* Elliptic curves and number-theoretic algorithms // Proc. International Congress of Mathematicians. Berkeley, 1986. 99–120.
4. *Lenstra H. W.* Factoring integers with elliptic curves // Ann. Math. Ser. 2. 1987. **126**, N 3. 649–673.

5. *Lenstra A. K., Lenstra H. W., Manasse M. S., Pollard J. M.* The number field sieve // Proc. 22nd ACM Symposium on Theory of Computing. Baltimore, 1990. 564–572.
6. *Charles D.* Computational aspects of modular forms and elliptic curves // PhD thesis. University of Wisconsin-Madison, 2005.
7. *Edixhoven B., Couveignes J.-M., de Jong R., Merkl F., Bosman J.* On the computation of coefficients of a modular form // <http://www.arxiv.org/abs/math.NT/0605244>. 2006.
8. *Knapp A.* Elliptic curves. Princeton: Princeton Univ. Press, 1992.
9. *Schoof R.* Elliptic curves over finite fields and the computation of square roots mod p // Math. Comp. 1985. **44**. 483–494.
10. *Canfield E., Erdős P., Pomerance C.* On a problem of Oppenheim concerning “factorisatio numerorum” // J. Number Theory. 1983. **17**. 1–28.